DEMONSTRATION OF STEREO VISION FOR DEORBIT DESCENT AND LANDING

David C. Sternberg,[*] **Timothy P. Setterfield,**[†] **Erik S. Bailey**[‡]**, Adnan I. Ansar**[§]**, and Andrew E. Johnson**[**]

Planetary landers need to reduce velocity at low altitude for soft landing. Traditionally, estimating velocity and altitude has been performed with radar sensors whose performance meets the specific mission needs. There are not very many options for these sensors and they are difficult to include in a flight system either due to obsolescence, prohibitive cost or difficulty in accommodation. Recently, alternative sensing modalities are being pursued including Doppler LiDAR and vision. This paper describes results from a recent helicopter field test of a binocular stereo vision system for deorbit descent and landing applications. The system consisted of two 18.6˚ field of view cameras mounted 1.7m apart. Post processing of the images showed ranging accuracy better than 1% up to 500m and 17 cm/s velocimetry accuracy at 37m. For a flight system these images could be input into an FPGA-based processor which processes dense stereo and visual odometry in less than 1 second to achieve the stereo ranging frame rates required for soft landing. When coupled with vision based Terrain Relative Navigation this stereo system enables landing accuracies on the order of 10m.

INTRODUCTION

Stereo vision is a common sensing modality for surface and near surface autonomous navigation. Range measurements are generated from image matching and triangulation between a single stereo pair while motion is estimated from image matches between stereo pairs. Stereo vision has been a critical component of autonomous rover navigation for the Mars Exploration Rovers[1] and the Mars Science Laboratory[2] missions. To enable safer and faster driving between scientific regions of interests, the Mars 2020 mission will use FPGA accelerated stereo vision[3] on the Vision Compute Element (VCE)[4]. Currently, image processing for dense stereo ranging and motion estimation (aka visual odometry) are computed in less than 1 second on the VCE.

Most lander missions to date have used radar to provide altitude and velocity measurements during landing. Radars are probably ideal sensors for this application, but they can be large and expensive or have

[*] Guidance and Control Engineer, Payload/Instrument Pointing Control Systems Engineering Group, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109

[†] Guidance and Control Engineer, GN&C Hardware and Testbed Development Group, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109

[‡] Group Supervisor, EDL Guidance and Control Systems Group, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109

[§] Group Supervisor, Aerial and Orbital Image Analysis Group, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109

[**] Principal Robotics Systems Engineer, Guidance and Control Section, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109

poor performance. Alternatives to radar, like Doppler LiDAR are worth looking into. Given the 1 Hz processing rate possible with the VCE, stereo vision is also an option for missions with slow descent (e.g. asteroid or comet lander missions) or moderate touchdown velocity requirements (e.g., Mars Phoenix[5] and InSight). If the mission already requires a VCE for Terrain Relative Navigation, then stereo vision can be added on cheaply and with low risk by adding two stereo cameras to the system.

The performance of stereo based navigation functions is limited by the range accuracy from the image matches which is governed by this simple equation

$$\Delta R = \frac{R^2 \Delta \theta}{B} \qquad (1)$$

where ΔR is the error in range, R is range, B is the separation or baseline between the stereo cameras and $\Delta \theta$ is the accuracy of the matches between images in radians (approximately the field of view / number of pixels × correlation accuracy). For rover navigation, the cameras are wide field of view, have a moderate number of pixels, and have narrow a baseline that must be accommodated on the rover mast or body. As shown in Table 1, the Mars 2020 NavCam stereo system has a range error of 14 cm at 10 m range.

Given the much greater velocities for landing applications, we need this level of range accuracy at higher altitude so that there is time correct for errors before landing. This can be achieved with a larger baseline, narrower field of view or more pixels in the image. As shown in Table 1, it is possible to obtain this 14 cm error at 85 m range by changing the baseline to 2 m, the field of view to 12˚, and using a camera with 2592 pixels across its field if view. Most landers have space for 2 m baselines; the challenge is to maintain knowledge of the cameras' orientation to the pixel level through dynamic events (e.g., thrusting, pyrotechnic deployments). Possible approaches include a rigid bar between the camera or augmenting the cameras with inertial sensors to measure the misalignment in real-time. Narrow field of view lenses are straightforward to implement. However, decreasing the field of view will reduce the allowable angular rate to ensure there is enough overlap in the stereo pairs used for visual odometry. Cameras with large format detectors are becoming common in space applications. The challenge for landing is receiving the image quickly after it is taken so processing can begin. The LCAM developed for the Mars 2020 Lander Vision System (LVS)[4] can provide a 1024×12024 image to the VCE in less than 100ms. The LCAM also has a raw detector size of 2592×2048 and with small changes to the LCAM processing this image could be output in 500ms. With windowing inside the camera this time could be dramatically reduced.

Table 1: Stereo Range Error Examples.

	Range (m)	Field of View (rad)	Number of Pixels (#)	Baseline (m)	3-σ Correlation Accuracy (px)	3-σ Range Error (m)
Mars 2020 Rover NavCam	10	1.5	1280	0.43	0.5	0.14
Proposed Stereo Vision System	85	0.2	2592	2.0	0.5	0.14
Field Test System	44	0.325	2592	1.71	1.0	0.14

The primary purpose of the Mars 2020 LVS is to provide a map relative localization (MRL) estimate for avoiding large hazards in the landing ellipse. For purely propulsive landers on airless bodies the LVS can be used for pin-point landing because retargeting will take only a small fraction of the fuel required for soft landing. If a map of sufficient resolution is available, position estimates better than 10 m can be obtained[6].

It is straightforward to modify the LVS to add stereo vision for altimetry and velocimetry. The VCE already has the processing capability for terrain relative navigation and stereo vision. A simple physical interface card for the stereo cameras would need to be added into the spare 3U slot in the VCE. The stereo cameras could be created from the LCAM electronics and a 12˚ FOV lenses. This would be in addition to the LCAM that is already part of the system. Some software and reprogrammable FPGA changes would also need to be made to interface to the stereo cameras and sequence the stereo processing after the MRL processing completes. With these straightforward and low risk modifications the LVS can become a complete sensor solution for soft 10 m level pin-point landing. The open question is whether, in practice, stereo vision altimetry and velocimetry will meet the predicted theoretical performance. As described below, a helicopter field test was conducted to answer this question.

OVERVIEW OF TEST CAMPAIGN

The purpose of the stereo vision helicopter field test campaign is to provide an early assessment of the functionality and performance of stereo vision for landing with images taken at relevant altitudes and over terrestrial analogs of planetary terrain. The system was flown on an ASTAR AS350B3E helicopter over a number of trajectories spanning the flight envelope within the safe operational capabilities of the helicopter over a variety of test sites. During each flight, the system captured images in real-time as would be expected during flight, though all data analysis was performed post-flight. Separate development work is aimed at implementing the algorithms used for the data analysis onboard the flight system.

EXPERIMENT COMPONENTS

A prototype stereo vision system (SVS) was built from commercial components, however the cameras used the same detector that is part of the Mars 2020 lander vision system[4]. The sensors were integrated with a laptop computer that served as a data repository and sensor controller.

Sensors: The core of the SVS system is a set of commercial-off-the-shelf (COTS) sensors. They include a pair of iDS UI-3180CP-M-GL 2592×2048 pixel CMOS monochrome cameras each with a Xenon-TOPAZ XN 2,0/38/0901 lens, and an xNAV 550 GPS system with two G5Ant-52AT1 GNSS antennas. All of these COTS components are suitable for use in these testing conditions because of both their rugged design and their heritage in other high vibration environments. The cameras were attached to the landing structure of the helicopter which resulted in the predicted stereo performance shown in the last row of Table 1.

Table 2 lists the sensor components. The G5 Ant GPS antennas were both connected to the xNAV 550, and their relative positions were measured using a Leica Total Station before testing so that the xNAV 550 could determine the correct position and orientation of the helicopter.

Table 2. SVS COTS Prototype Sensors.

Part No.	Description	Manufacturer
UI-3180CP-M-GL	Cameras (2) for acquiring stereo images	iDS
Xenon-TOPAZ XN 2,0/38/0901	Lens for each camera	Schneider
xNAV550	GPS receiver unit	Oxford Technical Solutions
G5Ant-52AT1	(2) GNSS Antennas, TNC Connector, Ground Plane	Oxford Technical Solutions
RT-Base S-2 NR	GPS Base Station	Oxford Technical Solutions

Compute Element: The sensors are connected to a compute element for data acquisition; the prototype SVS compute element is a Dell laptop. The laptop ran two vendor-supplied program suites for sensor data acquisition: StreamPix for the camera image streaming, saving, and camera settings management, and a combination of NAVconfig and NAVdisplay to set up and display, respectively, the output of the xNAV 550 GPS system. The laptop was also used for post-processing of the data.

Figure 1 shows a block diagram of the data connections between the sensors and the laptop computer. The GPS receiver detects the signals from overhead satellites, including a 1PPS signal. The xNav 550 unit uses the 1PPS signal for timing both when to save position and orientation estimates to internal memory and when to send an image acquisition trigger to the two cameras. The synchronicity of GPS data and image data allowed the two to be related during data analysis.

FIELD TESTING

The cameras and GPS antennas were mounted externally to the helicopter. The xNAV 550 and the laptop were located within the helicopter so that two operators could assess the status of the images (correcting the exposure time as necessary) and determine when the GPS position and attitude converged. The ability to monitor the state of the sensor systems enabled the operator to detect and correct errors in real time, reducing overall flight time while maximizing data returns.

Figure 2 shows an external view of the as-flown configuration of the externally-mounted sensors. The cameras were aligned so that their boresights were within 1˚ of each other, though verging occurred as a

result of the helicopter's weight lifting off the skids during flight. This verging was acceptable owing to the small change in relative boresights as compared to the camera fields of view. The GPS antennas were deliberately canted off-nadir by approximately 11.5˚ to the helicopter's starboard side to provide an improved view of overhead satellites.

The field testing occurred on September 21, 2018, over a period of two flights. The flights were spaced during the day to allow different sun angles to provide varying illumination conditions of the ground targets. The midday flight averaged approximately a 55˚ sun elevation angle, while the afternoon flight saw sun angles of approximately 35˚. This difference in sun angles not only changed the effect of shadows from the targets and terrain, but necessitated live changes by the laptop operator to the camera exposure parameters to prevent data loss in the acquired images
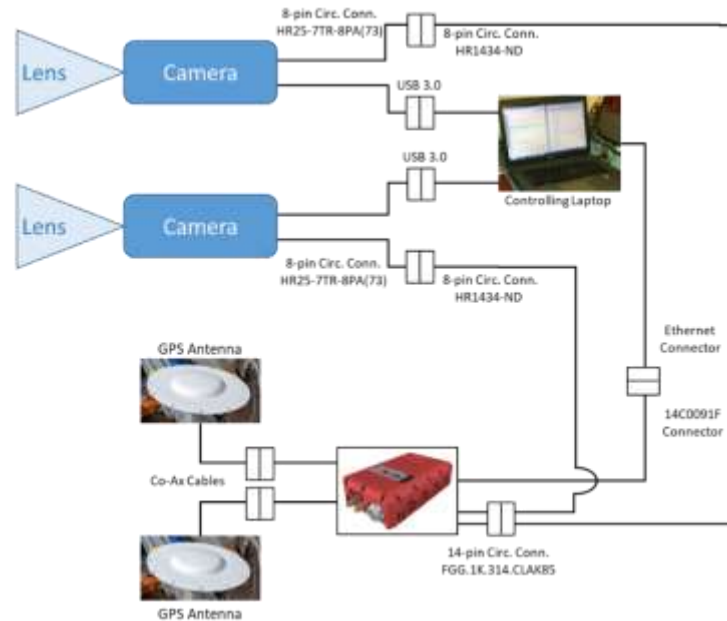
Figure 1. Prototype SVS System Block Diagram.

Figure 2. Photo of Sensors Mounted to AS350B3E.

Each flight originated at Barstow Airport (KDAG) and flew over four pre-defined target locations at the Pisgah lava flow in California. The ground crew were located near these locations to survey them and to provide guidance to help the helicopter identify its location relative to the targets, owing to the small field of view of the downward facing cameras. The four locations are shown in Figure 3: Transition East, Flat East, Dark East, and Rough East. These locations cover a wide range of albedo and surface roughness, allowing the field testing to acquire data over the representative set of expected planetary landing locations.

At each location were a set of surveyed targets designed to provide a ground truth for the cameras. Figure 4 (left) shows these calibration targets. The left target is a 2×2 array of 10" diameter circles on a 20" spacing, while the right target is a 5×5 array of 4" diameter targets on an 8" spacing. Figure 4 (right) shows a photo of the two calibration targets taken from within the helicopter. The two ground crew members are also visible near the top center of the picture for scale. This image was taken at approximately the lowest altitude above the ground that was flown during the testing; at lower altitudes, stirred dust degraded the view of the surface.

Figure 3. Target Locations for SVS Test Day (Google Earth).

Figure 4. (Left) Calibration Targets. (Right) View from Helicopter on Approach to Transition East Target Field.

In addition to the calibration targets, a set of surveyed hazard mounds were placed randomly in the Flat East target field over an area of about 50 m^2. These hemispheres were 24 and 30" across, providing a representative challenge for landing on another celestial body. The inclusion of these hazards in the flat testing area enables the images be re-used in future analyses aimed at hazard detection and hazard avoidance; herein, the mounds were used to visually assess the quality of dense depth maps. Figure 5 (right) shows a photo taken by the ground crew of the helicopter hovering over the hazard mounds at its lowest altitude.

GROUND TESTING

For ground testing, the same pair of iDS UI-3180CP-M-GL cameras were mounted 2 m apart on a rigid 80/20 extruded Aluminum bar with aligned boresights. This assembly was mounted on a tripod and

positioned atop the JPL parking structure overlooking the Arroyo Seco (see Figure 6). One team member on the ground faced the cameras with a textured patch and retroreflector prism. Two team members on top of the parking structure took stereo images and determined range using a Leica Total Station laser rangefinder. Images were taken of the textured patch at distances between 62 and 516 m at the approximate positions illustrated on the right hand side of Figure 6.

Figure 5. (Left) Hazard Mounds. (Right) Helicopter Hovering over Mounds at Flat East Target Field.

Figure 6. (Left) Ground Testing Setup. (Right) Approximate Positions of the Ground Testing Setup (Green) and Surveyed Locations (Red).

FLIGHT DATA ANALYSIS

Initial surveyed monocular CAHVOR camera calibrations[7] were performed on each camera using internal JPL calibration tools. Since the cameras were mounted to the helicopter struts, they were subject to vibrations at frequencies on the order of 10-20 Hz, driven by the main rotor. With a 1 Hz framerate, the relative motion (mainly rotation) between the two cameras precluded the use of a single, constant extrinsic calibration; to mitigate this issue, extrinsic calibration (i.e. determining the relative position and attitude of the cameras) was performed separately for each image pair using natural features. This approach required use of the GPS altitude and prevented the use of flight results to evaluate absolute depth accuracy. However, the flight results could be used to assess the precision of the range measurements and validate that stereo matching and triangulation can be used to generate ranges at high altitude under flight-like viewing conditions. Ground results, included in the next section, are used to assess the accuracy of depth measurements up to 500 m.

Calibration for the flight results involved initial and final intrinsic and extrinsic calibrations; an overview of the calibration procedure is illustrated in Figure 7.

Initial Intrinsic Calibration: To perform intrinsic camera calibration (i.e. determination of the focal length, skew, radial distortion, etc.), a 10×10 dot target pattern was imaged, where each dot is 2" in diameter and spaced 4" apart (center-to-center); a Leica Total Station laser rangefinder was used to obtain the range of the corner dots. Internal tools for surveyed calibration were used to determine intrinsic calibration from the monocular images.

Initial Extrinsic Calibration: To perform extrinsic camera calibration, the on-helicopter translation between the two cameras (baseline) was calculated by surveying three IR reflectors attached to each camera. This distance is believed to have been stable throughout the flight. The cameras' orientations, to which the resultant stereo results are very sensitive, were found to vary significantly throughout the flight. To calculate the cameras' orientation, natural SIFT features in the left and right images were detected and matched.
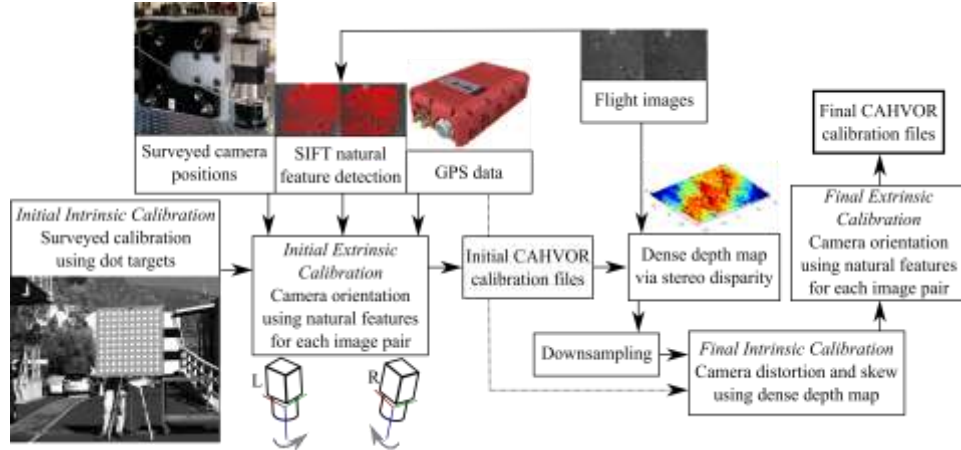
Figure 7. Flight Calibration Procedure.
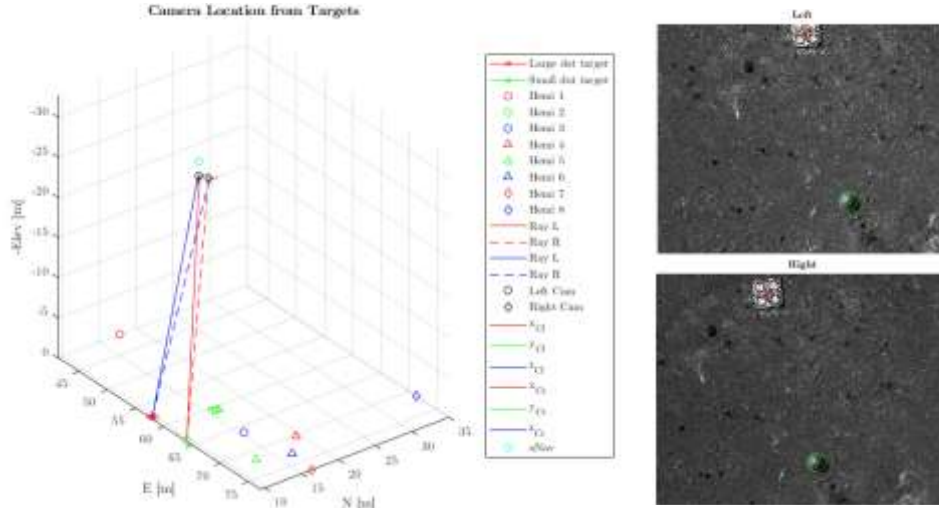
**Figure 8. Determining Initial Altitude through Ray Intersection.**

Flight data consisted of an ascent above surveyed ground targets, which were captured by both cameras. The initial, lowest altitude of the left camera was determined by calculating the direction vectors from two surveyed targets to the camera perspective center, and calculating their closest point of intersection (see Figure 8); the remaining camera altitudes were then calculated using the change in GPS altitude from this initial altitude. This method was used to remove a bias in GPS altitude. To determine the optimal camera orientations, angle-axis rotation vectors $\boldsymbol{a}_L, \boldsymbol{a}_R = \{\boldsymbol{a} \in \mathbb{R}^3 : |\boldsymbol{a}| \leq \pi\}$ were used to parameterize the orientation of the left and right cameras. Camera orientations that minimized the sum of squares of reprojection errors ρ_i and depth errors ϵ_i over N_p corresponding points were sought.

$$\boldsymbol{a}_L^*, \boldsymbol{a}_R^* = \operatorname{argmin} \sum_{i=1}^{N_p} \rho_i^2\big(\boldsymbol{a}_L, \boldsymbol{a}_R, \boldsymbol{p}_{L_i}, \boldsymbol{p}_{R_i}\big) + \epsilon_i^2\big(\boldsymbol{a}_L, \boldsymbol{a}_R, \boldsymbol{p}_{L_i}, \boldsymbol{p}_{R_i}, h_{C_l}, {}_{C_l}^W\boldsymbol{R}\big) \qquad (2)$$

Herein, reprojection errors ρ_i are determined by first finding the closest intersection, in 3D space, of rays passing through the SIFT feature point positions $\boldsymbol{p}_{L_i}, \boldsymbol{p}_{R_i}$ and the left and right camera perspective centers;

the location of these 3D points, in addition to camera orientation, depends on the camera intrinsic calibration and baseline between the cameras – both of which are kept constant. The generated 3D points are reprojected onto the left and right cameras, and the reprojection error ρ_i is the norm of the distance between the reprojected and original feature points in pixels. To compute depth errors ϵ_i, the imaged surface is assumed to be a level plane, with the left camera altitude h_{C_l} and relative orientation $^W_{C_l}\boldsymbol{R}$ determined by the xNAV550 GPS system. The norm of the perpendicular distance, in meters, between projected 3D points and the implied ground plane, scaled by $40/h_{C_l}$, is denoted depth error ϵ_i. The optimization in Equation (2) is performed several times in order to reject outliers. After each optimization, corresponding points with a reprojection error in excess of a prescribed tolerance are discarded, and N_p is reduced. At the final optimization step, the reprojection error tolerance for inliers is 2 pixels; in practice reprojection errors are typically < 0.5 pixels.

The left hand side of Figure 9 shows natural SIFT features for two image pairs with unmatched features in red, matched features in blue, and matched and filtered features in green (green results satisfy the epipolar constraint check and are used in extrinsic calibration). The right hand side of Figure 9 shows the vibration-induced change in angle-axis rotation between left and right cameras with respect to the first stereo pair.
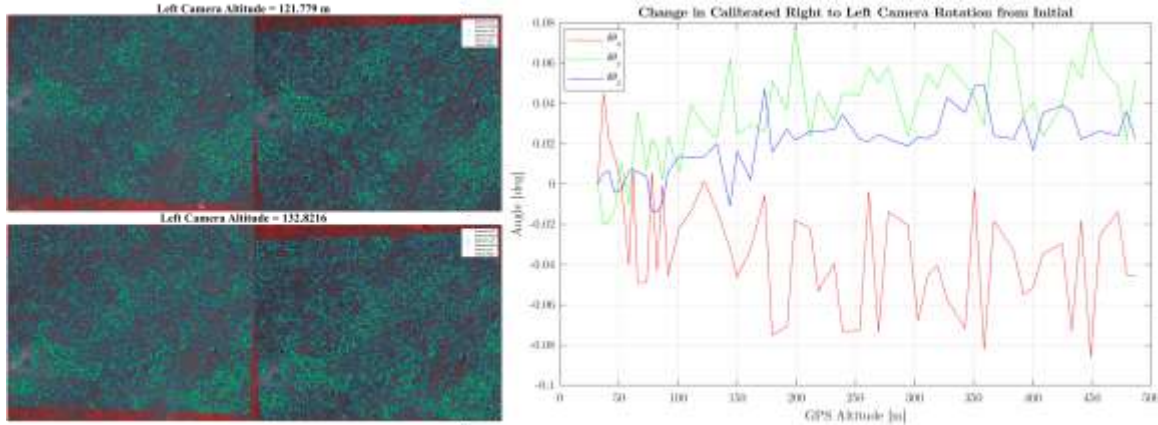
Figure 9. (Left) Image Pairs with SIFT Features. (Right) Right to Left Camera Rotation Change.
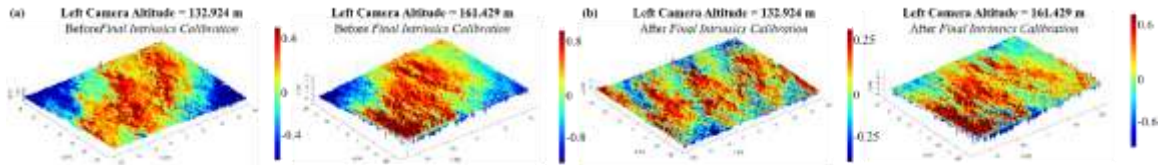
Final Intrinsic Calibration: Whereas the surveyed monocular calibration was performed at close distances on the order of 10 m, the flight stereo images were taken at distances of over 30 m. As a result, stereo-processed flight data is very sensitive to intrinsic calibration errors. Specifically, errors in knowledge of the non-perpendicularity between the horizontal and vertical camera axes were found herein to cause planarity deviations in the resultant depth maps. A final intrinsic calibration step was performed to address this concern. Dense depth maps from $m=21$ low altitude images were downsampled to reduce noise and accentuate non-planarity. Global values for CAHVOR optical axes $\boldsymbol{O}_L, \boldsymbol{O}_R \in S^2$, CAHVOR radial distortions $\boldsymbol{R}_L, \boldsymbol{R}_R \in \mathbb{R}^3$ and CAHVOR[7] skew terms $\boldsymbol{H}_L(2), \boldsymbol{V}_L(1), \boldsymbol{H}_R(2), \boldsymbol{V}_R(1) \in \mathbb{R}$ were sought that minimized the sum of squares of planarity deviation δ_i and reprojection error ρ_i over the m images (with N_d points each).

$$\boldsymbol{O}_L^*, \boldsymbol{O}_R^*, \boldsymbol{R}_L^*, \boldsymbol{R}_R^*, \boldsymbol{H}_L^*(2), \boldsymbol{V}_L^*(1), \boldsymbol{H}_R^*(2), \boldsymbol{V}_R^*(2) = \text{argmin} \sum_{i=1}^{mN_d} \delta_i^2(\boldsymbol{K}_L, \boldsymbol{K}_R, \boldsymbol{Z}_{1:m}) + \rho_i^2(\boldsymbol{K}_L, \boldsymbol{K}_R, \boldsymbol{Z}_{1:m}) \quad (3)$$

Above, $\boldsymbol{K}_L, \boldsymbol{K}_R$ are the full CAHVOR models including the modified extrinsics and candidate intrinsics, and $\boldsymbol{Z}_{1:m}$ are the m downsampled depth maps. To calculate the deviation from planarity δ_i, the norm of the perpendicular distance from a plane fit to each depth map was computed. The reprojection error ρ_i was calculated as in Equation (2); it was included in order to protect against large changes in intrinsics that could result in degenerate solutions.

Representative downsampled depth maps from $\boldsymbol{Z}_{1:m}$ are shown in Figure 10. While the imaged surface is nearly planar, the original depth maps are depressed at the top-left and bottom-right corners, and raised at the top-right and bottom-left corners. Following adjustment of the intrinsic calibration, the new results show the expected planarity.

Figure 10. Selected Depth Maps Before (a) and After (b) Final Intrinsic Calibration.

Final Extrinsic Calibration: The result of the final intrinsic calibration is then used to initialize a final pass of the SIFT feature-based extrinsic calibration procedure described above. The resultant image-specific calibrations are used for the remainder of the flight data image processing.

Flight Results: To predict the expected variance in depth measurements from the stereo camera system, an error model was constructed with the parameters in Table 3. Errors in depth caused by baseline stability grow linearly with depth, whereas those caused by correlation and misalignment grow quadratically (see Equation (1)), dominating the error model.

Table 3. Parameters in Stereo Error Model.

Stereo baseline	Camera horizontal field of view	Horizontal resolution	Correlation accuracy (3-σ)	Misalignment between cameras (3-σ)	Baseline stability (3-σ)
1.71 m	0.325 rad (18.6°)	2592 pixels	1 pixel	0.0003 rad (0.0172°)	0.002 m
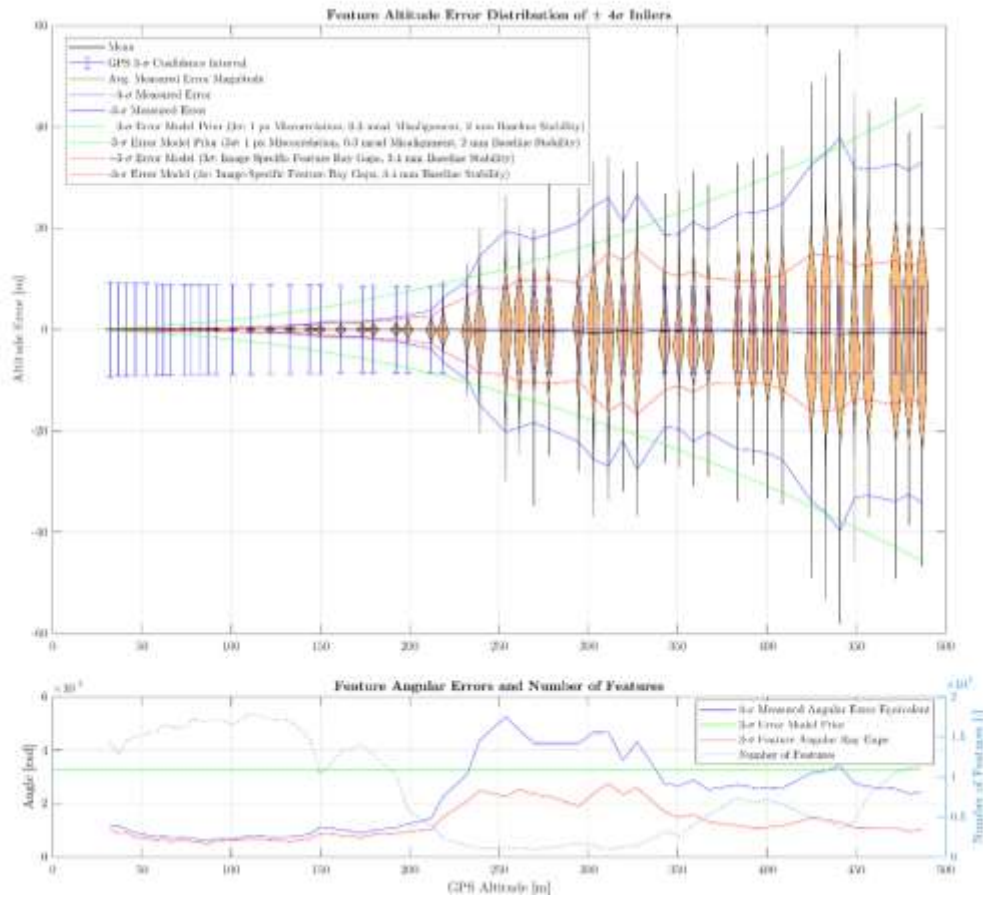
Figure 11. (Top) Flight Feature Altitude Error Distribution. (Bottom) Flight Feature Angular Error and Number of Features.

The errors in computed altitude of SIFT feature points, projected using the final CAHVOR calibrations, are used to compare the variance of the flight results against that of the 3-σ error model. For this comparison, all extreme altitude outliers falling outside 4-σ of the mean calculated altitude are removed. Figure 11 (top)

shows the distribution of the error in feature points' altitude compared with the model. The blue traces show the actual ±3-σ distribution of the features; the green traces show the modelled ±3-σ error given parameters from Table 3; finally, the red traces show the modelled quadratic error term using the computed feature angular ray gaps in each image. It is evident from the correlation of the red and blue traces that the measured variance of feature altitude originates predominantly from the features' angular ray gap. Generally, the blue traces stay within the bounds of the green traces, indicating that the parameters in Table 3 are conservative. However, between 239.1 m and 327.3 m, the measured error exceeds the model. This can be explained by Figure 11 (bottom), where: the blue trace shows the equivalent angular error of the ±3-σ altitude error; the green trace shows the modelled feature angular error from the combination misalignment and correlation in Table 3; and the red trace shows the actual angular ray gap. The dashed light-blue trace shows the number of inlying features used in extrinsic calibration. In the interval where the number of features fall, the angular feature error grows. The unexpectedly high error is therefore an artifact of the extrinsic calibration, present since the terrain imaged during this period was deficient in feature matches, resulting in a less accurate camera calibration and increased angular ray gap.

GROUND DATA ANALYSIS

Ground data is used to assess the absolute accuracy of long-range, wide-baseline stereo vision. Unlike the flight calibration procedure, the ground calibration procedure does not use any secondary measurement of depth to adjust the extrinsic calibration.

Calibration: In order to perform stereo calibration, 21 stereo image pairs were taken of the 10×10 dot target pattern at different depths and positions in the field of view. Camera intrinsics were initialized using the results from the *Initial Intrinsic Calibration*. An internal JPL unsurveyed stereo calibration tool was used to complete the calibration. Since the ground data was taken using cameras mounted on a rigid metal bar, it was not subject to vibration and a single camera calibration is used below for all images.

Ground Results: A dense depth map was obtained for each image using an internal JPL tool. To assess the accuracy of the results, the texture patch was first located in the image. The norm of distance between the left camera and the retroreflector prism, as measured by the Leica Total Station (LTS), was corrected for camera boresight angle to give an accurate measure of depth from the camera. Then, valid depths within the texture patch were extracted and compared with the measured depth. The resulting distribution of points and mean depth errors are shown in Table 4 and Figure 12. As the texture patch becomes more distant, fewer points are available to assess the statistics of the depth measurements. Below 500 m, the mean depth errors are all less than 1%, and the distribution of points stays within the 3-σ error model. At 516 m depth, the furthest distance for which measurements were taken, the results begin to degrade.
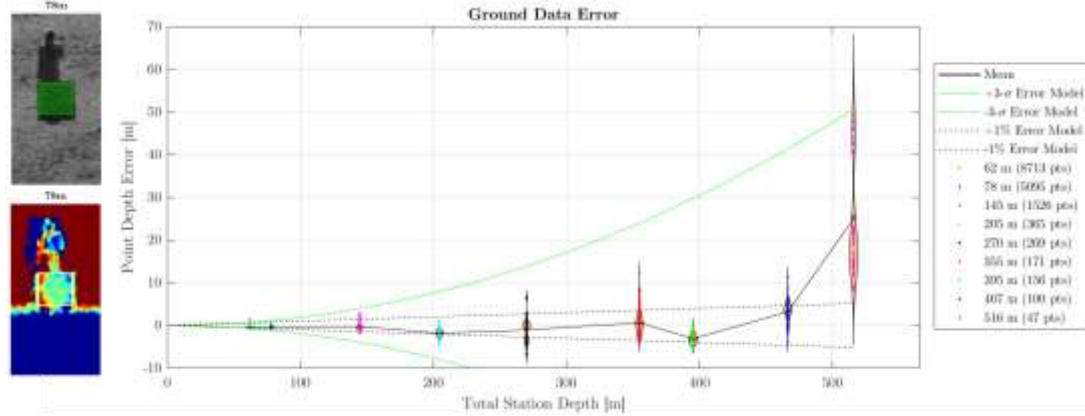
Table 4. Stereo Depth Error at Total Station Measured Depths.

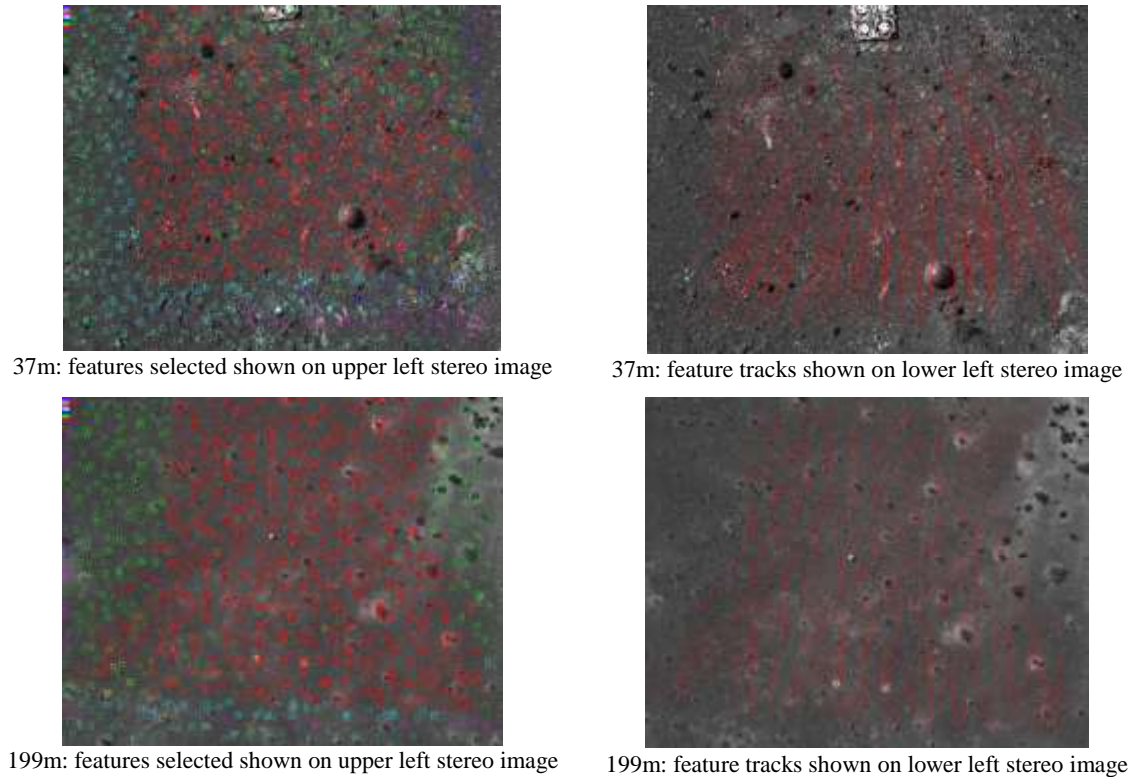| Measured Depth (LTS)	# of Points in Patch	Mean Stereo Depth Error	3-σ Stereo Depth Error
62.034 m	8713	-0.385 m (0.620 %)	0.372 m (0.599 %)
78.297 m	5095	-0.371 m (0.474 %)	0.417 m (0.533 %)
144.774 m	1526	-0.246 m (0.170 %)	1.568 m (1.083 %)
204.670 m	365	-1.848 m (0.903 %)	2.422 m (1.184 %)
270.170 m	269	-1.037 m (0.384 %)	7.016 m (2.597 %)
354.953 m	171	0.624 m (0.176 %)	7.745 m (2.182 %)
395.379 m	156	-3.097 m (0.783 %)	3.266 m (0.826 %)
466.546 m	100	3.181 m (0.682 %)	7.087 m (1.519 %)
515.977 m	47	25.023 m (4.850 %)	41.636 m (8.069 %)

STEREO VISUAL ODOMETRY PROCESSING

After the careful calibration and correction described above, the flight sequence was also used for visual odometry performance assessment. Consecutive stereo pairs from the sequence were processed by the MSL visual odometry software[3] which can be readily adapted to the VCE image processing pipeline. The images sequence between 200 m and 37 m altitude, binned down to 1024×1296 pixels, were processed using the same set of parameters and two examples are given in Figure 13. The top row shows the result of matching

stereo pairs taken at 37 m and 32 m altitude. The left image in the figure shows the features selected for matching between all 4 images. These features run through multiple processing steps (matching, triangulation, outlier rejection etc.); the ones that survive to be used for motion estimation are shown in red superimposed on the higher altitude image (i.e., 37 m). The right image in the figure shows the tracks between the two left images in the stereo pairs displayed on the lower altitude image (i.e., 32 m). The second row of the figure shows the same type of result for an image pair taken with images at 199 and 192 m altitude.


Figure 12. (Left) Patch in Image and Depth Map at 78 m. (Right) Distribution of Ground Data Error.

37m: features selected shown on upper left stereo image

37m: feature tracks shown on lower left stereo image

199m: features selected shown on upper left stereo image

199m: feature tracks shown on lower left stereo image

Figure 13. Examples of Selected Features (Left) and Tracks from Image Correlation (Right) for Visual Odometry Results at 37 m (Top) and 199 m (Bottom).

Each set of consecutive stereo pairs generated a visual odometry estimate that can be compared to the GPS derived ground truth described above. The errors, which combine errors in the visual odometry estimates and the ground truth, are shown in Figure 14. Even though the GPS ground truth position has a predicted precision of ±6 m 3-σ, the errors are much smaller than this. The errors are expected to decrease as the altitude

decreases and the results show this trend. The errors at 37 m altitude are (0.08, 0.12, 0.09) m which is small enough for soft touchdown for most planetary landing missions. The errors grow as altitude increases but in a landing mission these errors can be gradually eliminated during descent using the powered descent control system. The visual odometry result at 132 m altitude is out of family with the rest of the results. Because this result has many feature tracks, this large error is likely due to errors in the camera extrinsic or intrinsic calibration described above for one of the stereo pairs and not due to the visual odometry processing.
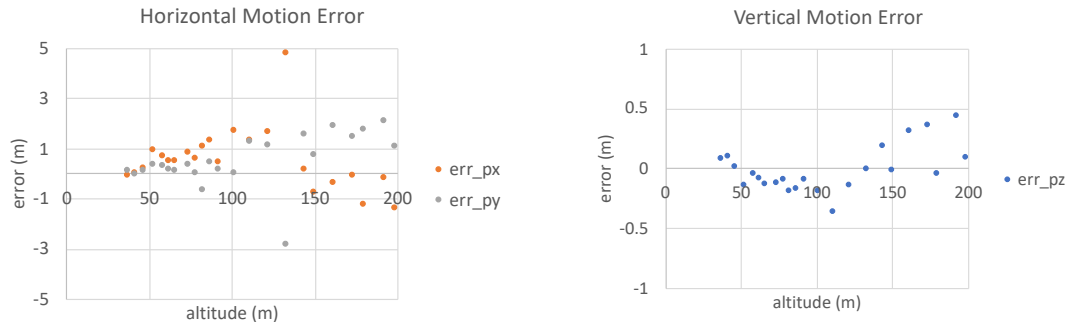
Figure 14. Visual Odometry Processing of Flight Sequence. (Left) Horizontal Errors and (Right) Vertical Errors.

CONCLUSION

The applicability of stereo vision as a sensing modality for altimetry and velocimetry for planetary landing applications involving slow descent is investigated herein. A helicopter field test is used in combination with ground tests to validate theoretical performance of the stereo camera. The single feature altitude measurement variance was lower than the error model in flight test results, except where feature-poor regions precluded accurate ad-hoc extrinsic calibration. Better than 1% range accuracy was demonstrated through ground testing for ranges up to 467 m. At the lowest field-test altitudes of 37 and 32 m, 1 Hz visual odometry accuracy implies a single-measurement horizontal velocity error of 0.14 m/s, and a vertical velocity error of 0.09 m/s, which are low enough for soft touchdown in most planetary landing missions.

ACKNOWLEDGMENTS

REFERENCES

[1] M. Maimone, Y. Cheng and L. Matthies, "Two Years of Visual Odometry on the Mars Exploration Rovers." *Jour. Field Robotics: Special Issue on Space Robotics.* Vol 24, No.4, pp. 169-186, March 2007.

[2] A. Johnson, S. Goldberg, Y. Cheng and L. Matthies "Robust and Efficient Stereo Feature Tracking for Visual Odometry." *Proc. IEEE International Conference on Robotics and Automation (ICRA08).* May 2008.

[3] T. Howard, A. Morfopolous, J. Morrison, Y. Kuwata, C. Villalpando, L. Matthies and M. McHenry," Enabling continuous planetary rover navigation through FPGA stereo and visual odometry." *Proc IEEE Aerospace Conference (IEEEAERO2012).* April 2012.

[4] A. Johnson S. Aaron, Johnny Chang, Y. Cheng, J. Montgomery, S. Mohan, S. Schroeder, B. Tweddle, N. Trawny and J. Zheng, "The Lander Vision System for Mars 2020 Entry Descent and Landing." *Proc. AAS Guidance Navigation and Control Conference (AAS-17-038).* February 2017.

[5] P. Desai, J. Prince, E. Queen, J. Cruz and R. Grover, "Entry, Descent, and Landing Performance of the Mars Phoenix Lander." *AIAA Journal Spacecraft and Rockets.* Vol 48, No. 5, Sept-Oct 2011.

[6] N. Trawny et al., "Flight testing of terrain-relative navigation and large-divert guidance on a VTVL rocket," *Proc. AIAA Space Conference*, 2015.

[7] K. Di and R. Li, "CAHVOR camera model and its photogrammetric conversion for planetary applications." *Journal of Geophysical Research E: Planets.* Vol. 109, No. 4, 2004.